Федеральное государственное автономное образовательное

учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.01 «Программная инженерия» –

Информатика и вычислительная техника

**Отчёт**

**По УИР №1**

**«Обработка результатов измерений:**

**статистический анализ числовой последовательности»**

**По моделированию**

**Вариант: 52**

**Выполнил:**

студент 3 курса  
Антипин Г. В.

**Группа:** P3332

**Преподаватель:**

Алиев Тауфик Измайлович

г. Санкт-Петербург, 2025

# Задание

# Цель работы

# Изучение методов обработки и статистического анализа результатов измерений на примере заданной числовой последовательности путем оценки числовых моментов и выявления свойств последовательности на основе корреляционного анализа, а также аппроксимация закона распределения заданной последовательности по двум числовым моментам случайной величины.

**Содержание отчета**

1. оценки математичес*кого ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации* заданной числовой последовательности и *доверительные интервалы* для оценки математического ожидания с доверительными вероятностями 0,9; 0,95 и 0,99, сведенные в таблицу (форма 1);
2. график (график 1) значений заданной числовой последовательности с результатами анализа характера числовой последовательности (возрастающая, убывающая, периодичная и т.п.);
3. результаты автокорреляционного анализа (значения коэффициентов автокорреляции со сдвигом 1, 2, 3, …), представленные как в числовом (форма 3), так и графическом виде, с *обоснованным выводом о характере заданной числовой последовательности* (можно ли ее считать случайной);
4. гистограмма распределения частот для заданной числовой последовательности (график 2);
5. параметры, рассчитанные по двум начальным моментам и определяющие *вид аппроксимирующего закона распределения* заданной случайной последовательности (равномерный; экспоненциальный; нормированный Эрланга; гипоэкспоненциальный; гиперэкспоненциальный);
6. *описание алгоритма (программы) формирования* аппроксимирующего закона распределения и расчета значений всех числовых характеристик с иллюстрацией (при защите отчета) его работоспособности;
7. выводы по результатам сравнения сгенерированной в соответствии с полученным аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности, а именно:
   1. сравнения *плотности распределения* аппроксимирующего закона с *гистограммой распределения* частот для исходной числовой последовательности (график 3);
   2. расчета числовых характеристик *сгенерированной* в соответствии с аппроксимирующим законом распределения случайной последовательности: математического ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации (представленные в таблице по форме 2) и коэффициентов автокорреляции при разных значениях сдвигов (в таблице по форме 3), а также сравнения (в %) полученных значений со значениями, рассчитанными для *заданной* числовой последовательности;
   3. проведения *корреляционного анализа* сгенерированной в соответствии с аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности на основе *коэффициента корреляции*.
8. ***по каждому из перечисленных выше пунктов отчета должны быть сформулированы результативные выводы и заключения****.*

# Ход работы

**Этап 1.** Оценки *математического ожидания*, *дисперсии*, *среднеквадратического отклонения*, *коэффициента вариации* заданной числовой последовательности и *доверительные интервалы* для оценки математического ожидания с доверительными вероятностями 0,9; 0,95 и 0,99, сведенные в таблицу.

Характеристики заданной ЧП (вариант 52)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Характеристика** |  | **Количество случайных величин** | | | | | |
|  | **10** | **20** | **50** | **100** | **200** | **300** |
| **Мат. ож.** | Знач. | 166.72 | 176.09 | 164.13 | 167.24 | 175.91 | 183.19 |
| % | -8.99 | -3.88 | -10.40 | -8.71 | -3.97 |
| Дов. инт. (0,9) | Знач. | ±56.03 | ±37.97 | ±22.75 | ±19.24 | ±13.13 | ±10.95 |
| % | 411.80 | 246.83 | 107.80 | 75.78 | 19.97 |
| Дов. инт. (0,95) | Знач. | ±65.14 | ±45.96 | ±27.27 | ±23.00 | ±15.67 | ±13.06 |
| % | 429.53 | 252.00 | 108.84 | 76.14 | 20.00 |
| Дов. инт. (0,99) | Знач. | ±99.35 | ±62.83 | ±36.35 | ±30.44 | ±20.67 | ±17.18 |
| % | 478.33 | 265.75 | 111.61 | 77.22 | 20.31 |
| **Дисперсия** | Знач. | 9343.88 | 9635.16 | 9212.08 | 13438.91 | 12627.25 | 13206.82 |
| % | -29.25 | -26.97 | -30.25 | +1.76 | -4.39 |
| **С. к. о.** | Знач. | 96.66 | 98.21 | 95.98 | 115.93 | 112.37 | 114.92 |
| % | -15.89 | -14.54 | -16.48 | -0.87 | -2.22 |
| **К-т вариации** | Знач. | 0.580 | 0.558 | 0.585 | 0.693 | 0.639 | 0.627 |
| % | -7.58 | -11.09 | -6.79 | 10.50 | 1.82 |

Вывод:

- Математическое ожидание на малых объёмах выборки изменяется нерегулярно, но при увеличении количества данных постепенно стабилизируется и к N = 300 достигает 183,19, что говорит о сходимости оценки.

- Доверительные интервалы заметно сужаются с ростом выборки: если при N = 10 разброс велик, то начиная с N = 200 ширина интервала отличается от итоговой менее чем на 20 %, что указывает на повышение точности.

- Дисперсия и среднеквадратичное отклонение после N = 100 практически не меняются, что свидетельствует о стабилизации разброса данных.

- Коэффициент вариации достигает максимума при N = 100 (0,693) и затем снижается до 0,627 при N=300, что говорит об умеренной вариабельности и устойчивости распределения.

- Итог: оценки становятся надёжными при N ≥ 200 при меньших объёмах наблюдается повышенная нестабильность результатов.

**Этап 2.** Значений заданной ЧП с результатами анализа характера числовой последовательности.

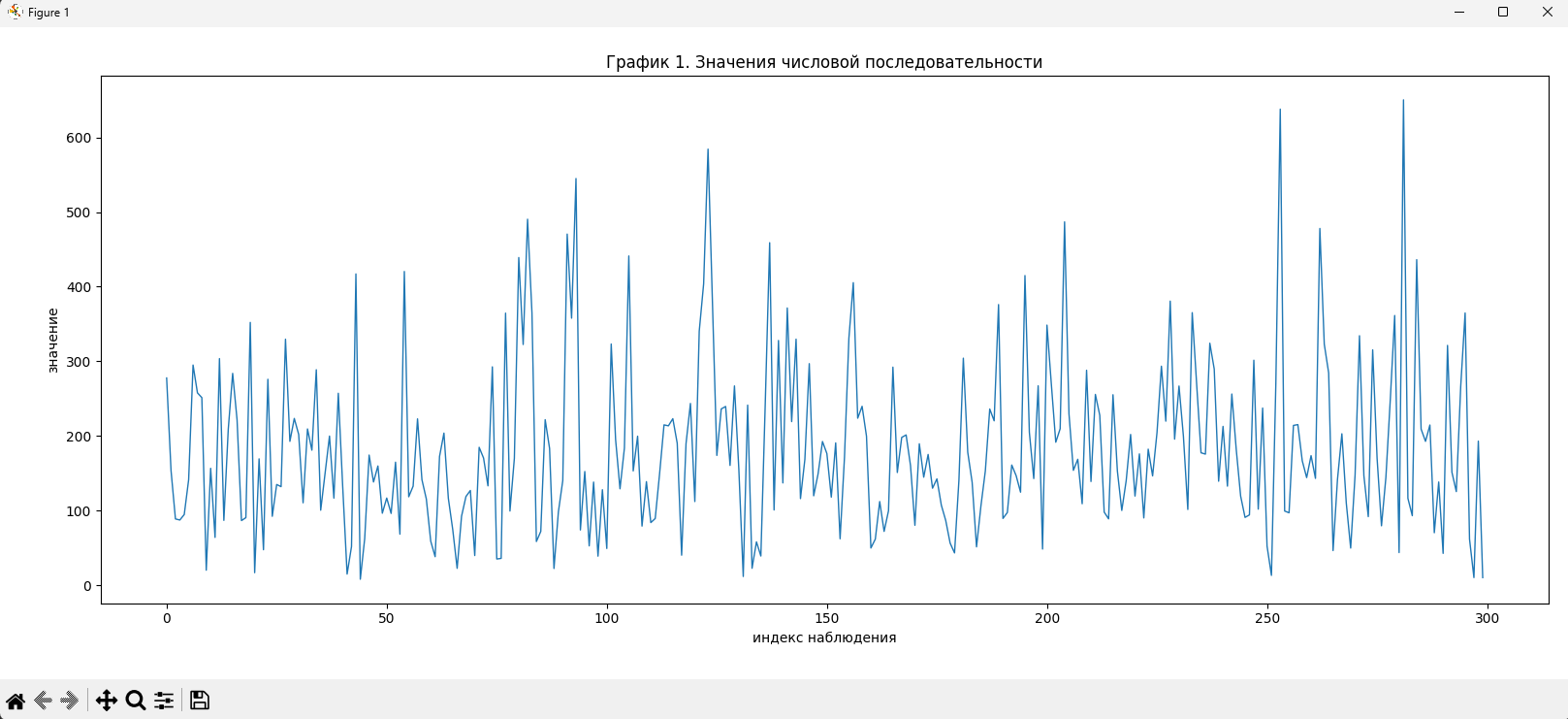


График №1

**Вывод:** Последовательность не имеет ярко выраженного тренда — значения хаотично колеблются, отсутствует устойчивое возрастание или убывание, наблюдаются отдельные пики и провалы, но они распределены неравномерно — то есть периодичность отсутствует

**Этап 3.** Результаты автокорреляционного анализа (значения коэффициентов автокорреляции со сдвигом 1, 2, 3, …), представленные как в числовом (форма 3), так и графическом виде.

Коэффициенты автокорреляции

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Сдвиг ЧП** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **К-т АК** | 0.090 | 0.035 | 0.024 | 0.003 | -0.079 | -0.024 | -0.050 | -0.050 | 0.103 | 0.061 |

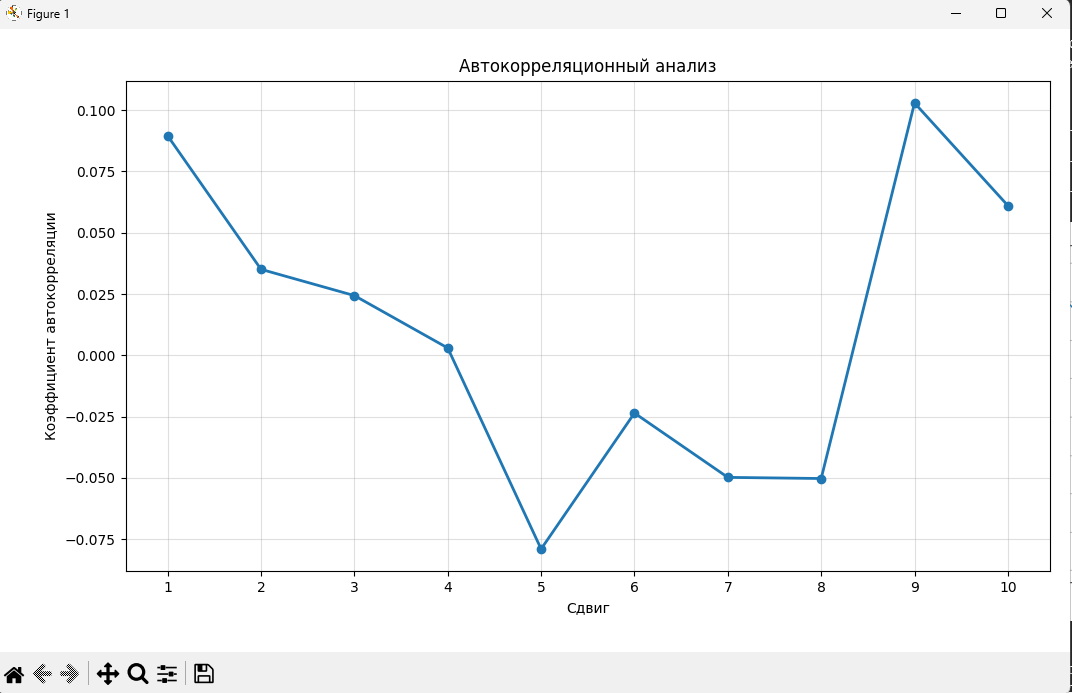
****

График №2

**Вывод:** Последовательность можно считать случайной, так как автокорреляционный анализ не выявляет существенной зависимости: коэффициенты АКФ малы и колеблются около нуля, все они попадают в ориентировочные границы значимости. Устойчивого убывающего/осциллирующего рисунка или выраженных пиков нет, периодичность не наблюдается.

**Этап 4.** Гистограмма распределения частот для заданной числовой последовательности (график 2).

Интервалы гистограммы

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Интервалы** | | | | | | | | | | | | | | | | | | |
| **№** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| **Левая граница** | 7.39 | 43.62 | 79.31 | 115.01 | 150.70 | 186.39 | 222.08 | 257.77 | 293.46 | 329.16 | 364.85 | 400.54 | 436.23 | 471.92 | 507.61 | 543.30 | 579.00 | 614.69 |
| **Правая граница** | 43.62 | 79.31 | 115.01 | 150.70 | 186.39 | 222.08 | 257.77 | 293.46 | 329.16 | 364.85 | 400.54 | 436.23 | 471.92 | 507.61 | 543.30 | 579.00 | 614.69 | 650.38 |
| **Частота** | 20 | 27 | 41 | 51 | 39 | 38 | 21 | 17 | 12 | 12 | 5 | 5 | 5 | 3 | 0 | 1 | 1 | 2 |

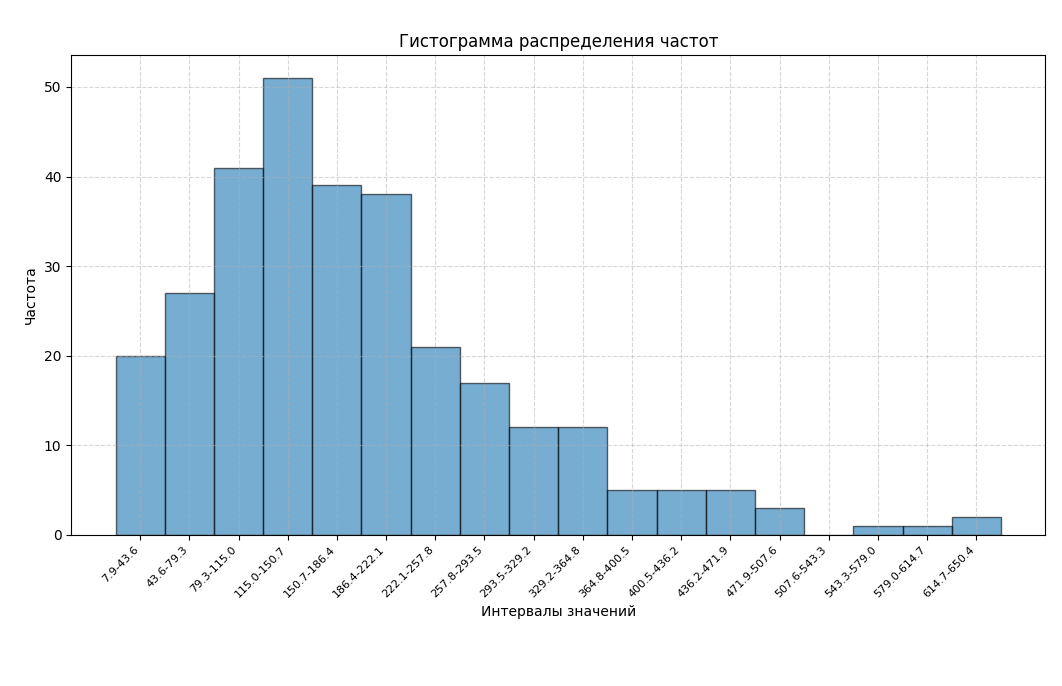


График №3

Вывод: Исходя из построенной гистограммы, можно предположить, что закон распределения числовой последовательности близок к закону Эрланга.

Большая часть значений числовой последовательности сосредоточена в диапазоне примерно [8; 330], где наблюдаются наибольшие частоты, тогда как оставшаяся часть диапазона образует длинный правый «хвост», содержащий редкие крупные значения до 650.

**5 этап.** Параметры, рассчитанные по двум начальным моментам и определяющие *вид аппроксимирующего закона распределения* заданной случайной последовательности (равномерный; экспоненциальный; нормированный Эрланга; гипоэкспоненциальный; гиперэкспоненциальный).

Коэффициент вариации меньше единицы, поэтому для аппроксимации последовательности будем использовать закон Эрланга k-го порядка.

Функция плотности вероятности для распределения Эрланга k-го порядка имеет вид:

Найдем параметры плотности распределения по выборке из 300 элементов.

*– порядок распределения*

*– параметр масштаба*

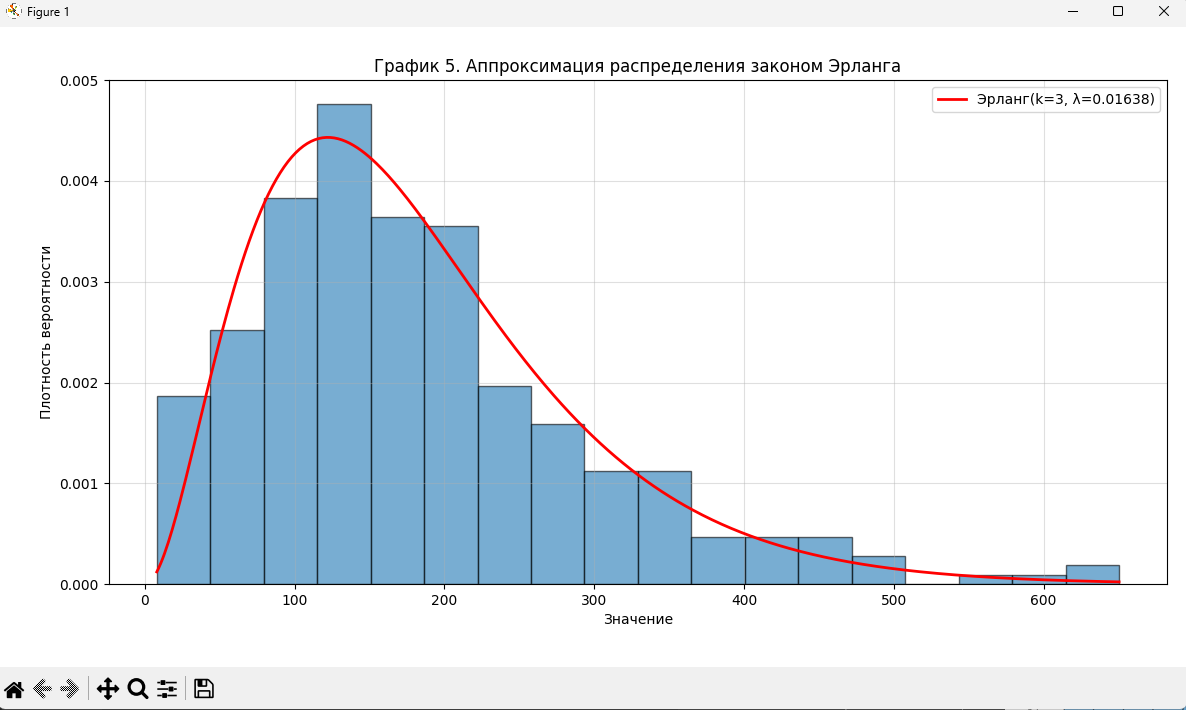


График №4

**Вывод:** Аппроксимация показала, что закон распределения числовой последовательности близок к закону Эрланга 3-го порядка с параметром λ = 0,01638. Теоретическая кривая хорошо описывает основную часть гистограммы — область наибольших частот до примерно 330 единиц. Незначительные расхождения в правом «хвосте» распределения объясняются наличием редких крупных значений и конечным объёмом выборки. В целом аппроксимация считается удовлетворительной, а выбранный закон Эрланга — адекватной моделью распределения для данной последовательности.

**6 этап.** *Описание алгоритма (программы) формирования* аппроксимирующего закона распределения и расчета значений всех числовых характеристик с иллюстрацией (при защите отчета) его работоспособности.

Для реализации распределения Эрланга использовался язык программирования Python.

|  |
| --- |
| import random  import math  def erlang\_generator(k, l):  result = 0.0  for i in range(k):  u = random.uniform(0, 1)  exp = - (1/l) \* math.log(u)  result += exp  return result  sample\_size = 300  k = 3  l = 0.01638  erlang\_samples = [erlang\_generator(k, l) for \_ in range(sample\_size)] |

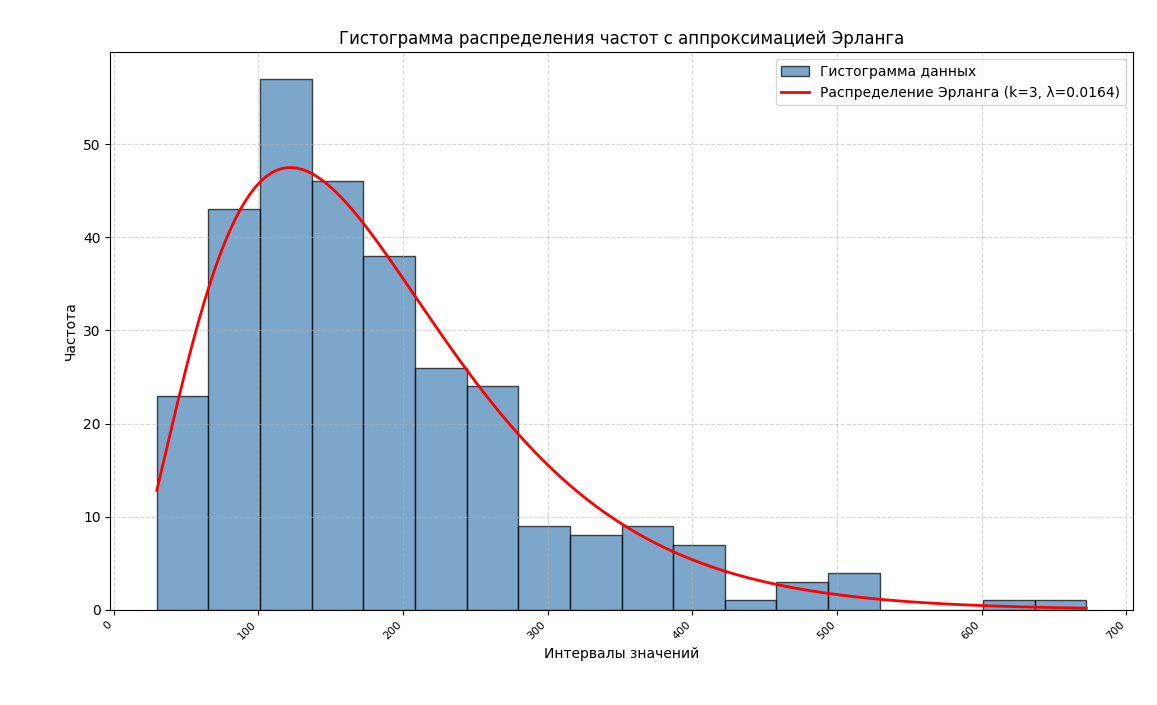


График №5

**Вывод:** Алгоритм генерации величин по закону Эрланга был реализован. Гистограмма и истинное распределение Эрланга показывают, что алгоритм был реализован правильно.

**7 этап.** Выводы по результатам сравнения сгенерированной в соответствии с полученным аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности.

Характеристики сгенерированной случайной ЧП

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***Закон распределения: Эрланг 3го порядка*** | | | | | | | |
| **Характеристика** |  | **Количество случайных величин** | | | | | |
|  | **10** | **20** | **50** | **100** | **200** | **300** |
| **Мат. ож.** | Знач. | 235.236 | 199.755 | 200.587 | 203.938 | 198.857 | 194.128 |
| % | 41.096 | 13.439 | 22.212 | 21.943 | 13.044 | 5.970 |
| Дов. инт. (0,9) | Знач. | ±56.856 | ±35.304 | ±27.163 | ±17.748 | ±12.672 | ±10.912 |
| % | 1.474 | −7.02 | +19.40 | −7.75 | −3.49 | −0.35 |
| Дов. инт. (0,95) | Знач. | ±67.748 | ±42.068 | ±32.367 | ±21.149 | ±15.1 | ±13.003 |
| % | +4.00 | −8.47 | +18.69 | −8.05 | −3.64 | −0.44 |
| Дов. инт. (0,99) | Знач. | ± 89.035 | ± 55.287 | ± 42.538 | ± 27.794 | ±19.844 | ±17.089 |
| % | −10.38 | −12.01 | +17.02 | −8.69 | −4.00 | −0.53 |
| **Дисперсия** | Знач. | 11947.904 | 9213.718 | 13635.864 | 11643.041 | 11870.327 | 13203.916 |
| % | +27.87 | −4.37 | +48.02 | −13.36 | −5.99 | −0.02 |
| **С. к. о.** | Знач. | 109.306 | 95.988 | 116.773 | 107.903 | 108.951 | 114.908 |
| % | +13.08 | −2.26 | +21.66 | −6.92 | −3.04 | −0.01 |
| **К-т вариации** | Знач. | 0.465 | 0.481 | 0.582 | 0.529 | 0.548 | 0.592 |
| % | −19.83 | −13.80 | −0.51 | −23.67 | −14.24 | −5.58 |

**Вывод:** Вычисленные характеристики сгенерированной ЧП показали небольшие отклонения от соответствующих исходных данных. Это показывает, что модель распределения Эрланга точно воспроизводит статистические свойства исходной последовательности.

Коэффициенты автокорреляции исходной и сгенерированной ЧП

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Сдвиг ЧП** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **К-т АК** исход. ЧП | 0.090 | 0.035 | 0.024 | 0.003 | -0.079 | -0.024 | -0.050 | -0.050 | 0.103 | 0.061 |
| **К-т АК** сгенерир. ЧП | -0.013 | -0.028 | 0.05 | -0.002 | -0.001 | -0.058 | -0.133 | 0.025 | 0.051 | 0.076 |
| **%** | -114.242 | -178.583 | 109.907 | -168.077 | -98.786 | 139.673 | 166.51 | -150.941 | -50.58 | 24.39 |

**Вывод:** для моей выборки коэффициенты автокорреляции на лагах 1–10 малы по модулю (все < 0.15), меняют знак и существенно колеблются от лага к лагу. При размере выборки ~300 их абсолютные значения в основном лежат внутри 95% доверительных границ (≈±0.11); лишь на 7 сдвиге наблюдается небольшое отклонение (≈−0.133), которое всё равно укладывается в 99% пределы (≈±0.15). Статистически значимой автокорреляции не выявлено - последовательность ведёт себя как случайная, аналогично исходной.

**Коэффициент корреляции между двумя числовыми последовательностями:**

**Вывод: д**ля двух исследуемых числовых последовательностей рассчитан коэффициент корреляции Пирсона, значение которого оказалось близким к нулю (r ≈ 0.047). Это указывает на крайне слабую линейную зависимость между выборками. Таким образом, изменение значений одной последовательности практически не влияет на другую, что подтверждает отсутствие выраженной линейной связи и согласуется с тем, что основное внимание уделялось совпадению распределений, а не их корреляции.

Значения сгенерированной ЧП

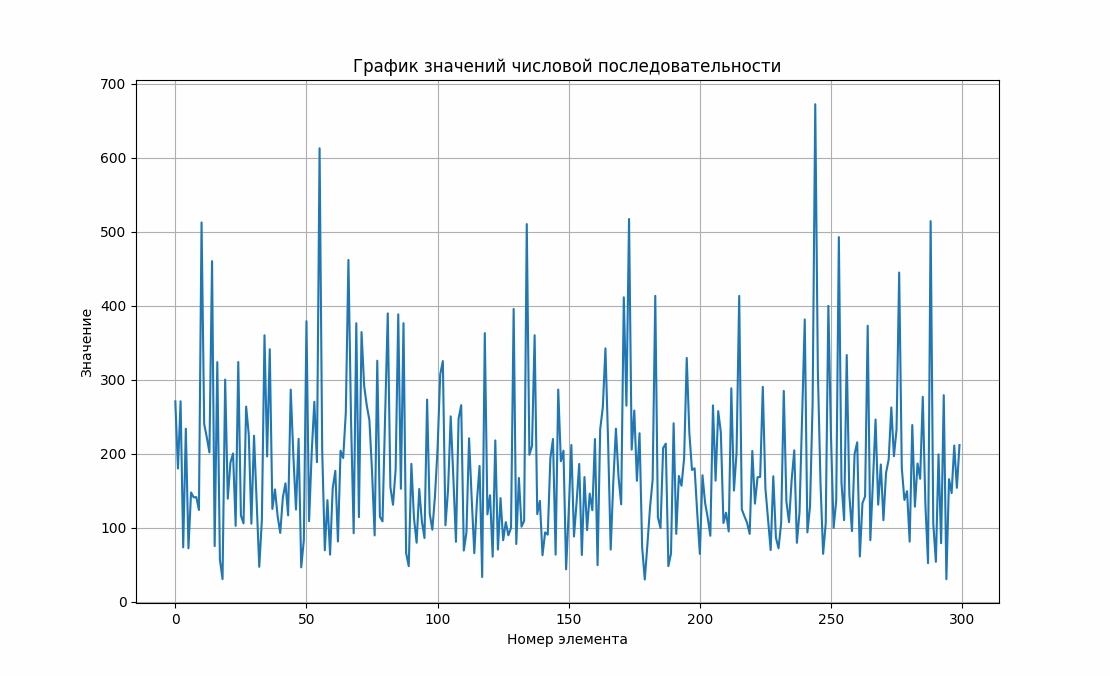
****

График №6

**Вывод:** Сгенерированная числовая последовательность, полученная по аппроксимирующему закону Эрланга 3-го порядка, внешне отличается от исходной последовательности. Это объясняется тем, что метод Эрланга моделирует лишь распределение значений, но не воспроизводит временную или порядковую структуру исходных данных. Таким образом, форма графика отражает статистические свойства распределения, а не последовательность изменений исходного сигнала.

**Выводы**

* Проведённый анализ аппроксимации показал, что распределение Эрланга третьего порядка хорошо описывает исходные данные. Это подтверждается близостью гистограмм частот для исходной и сгенерированной выборок.
* Статистические параметры сгенерированной последовательности лишь незначительно отличаются от исходных, что свидетельствует о высокой точности модели распределения Эрланга при передаче основных свойств исходных данных.
* Таким образом, результаты аппроксимации и сравнения исходной и смоделированной последовательностей подтверждают высокое соответствие как по визуальному виду, так и по числовым характеристикам.